Explainable NLP Model for Predicting Patient Admissions at Emergency Department Using Triage Notes

Emilien Arnaud [1,2], Mahmoud Elbattah [2,3], Pedro A. Moreno-Sánchez [4], Gilles Dequen [2], Daniel Aiham Ghazali [1]
[1] Emergency Department, Amiens-Picardy University, Amiens, France
[2] Laboratoire MIS, Université de Picardie Jules Verne, Amiens, France
[3] College of Art, Technology and Environment, University of the West of England, Bristol, United Kingdom
[4] Faculty of Medicine and Health Technology, Tampere University, Finland
arnaud.emilien@chu-amiens.fr, mahmoud.elbattah@u-picardie.fr

***Abstract*— Explainable Artificial Intelligence (XAI) has the potential to revolutionize healthcare by providing more transparent, trustworthy, and understandable predictions made by AI models. To this end, the present study aims to develop an explainable NLP model for predicting patient admissions to the emergency department based on triage notes. We utilize transformer models to leverage the extensive textual data captured in triage notes, while also delivering interpretable results by using the LIME approach. The results show that the proposed model provides satisfactory accuracy along with an interpretable understanding of the factors contributing to patient admission. In general, this work highlights the potential of NLP in improving patient care and decision-making in emergency medicine.**

Keywords— Explainable Artificial Intelligence (XAI), Natural Language Processing, BERT, Transformers, Emergency Department, Triage.

I. Introduction

The Natural Language Processing (NLP) field has seen rapid developments in recent years, with new models and techniques being introduced regularly. One of the key advancements in NLP has been the introduction of transformers [1], a type of neural network architecture that has become the standard for many NLP tasks. The transformer architecture was first introduced in 2017 and has since been adopted and adapted by researchers and practitioners for a variety of NLP applications, including sentiment analysis [2], machine translation [3], and question answering [4], among others. The success of transformers can be attributed to their ability to process long sequences of text, their parallelizable structure, and their ability to handle large amounts of data.

In the healthcare domain, the use of transformer models provides an opportunity to make better use of vast amounts of unstructured data, such as clinical notes, to support decision-making and improve care outcomes [5]. However, the adoption of these models also raises important ethical and interpretability concerns, as medical decisions can have significant consequences for patients and it is important to understand the reasoning behind these decisions. As such, the development of explainable AI (XAI) techniques for transformer models is an important area of research in the healthcare domain.

Despite the potential benefits of XAI in healthcare, such as increasing patient trust in AI systems, there has been a lack of studies that have applied XAI for transformers in this domain. This is likely due to the technical complexity of XAI, as well as the regulatory and ethical considerations that come with handling sensitive medical data [6]. As AI continues to play an increasingly important role in healthcare, it is crucial for further efforts to develop and apply XAI, to ensure that AI systems are both effective and trustworthy.

The present study aims to investigate the potential of explainable NLP methods within the context of an emergency department. Our goal is to leverage state-of-the-art NLP models, specifically BERT-based models [7], for accurate patient hospitalization predictions. To ensure transparency and interpretability, we incorporate XAI techniques that highlight the crucial words or phrases in the input text considered by the model during the prediction process. By offering a clear understanding of the decision-making process, this approach aims to foster trust and confidence in AI. It is important to note that our primary focus lies in exploring the explainability aspect of NLP models, rather than solely prioritizing higher performance compared to our earlier work [8].

II. Related Work

The application of XAI in healthcare can be seen in a variety of areas, including but not limited to, medical imaging, patient diagnosis, drug discovery, and clinical decision support systems. This section aims to examine various applications of XAI in the healthcare field, through a review of notable works that have put these methods into practice.

Numerous studies in the literature have explored the use of XAI for medical imaging data, including Computed Tomography (CT) scans, Magnetic Resonance Imaging (MRI) scans, and X-Rays. For instance, recent developments (e.g. [9-12]) have been focused on using XAI to help physicians and clinicians better analyze chest X-Ray images of patients infected with COVID-19. GradCAM [13], or the Gradient-weighted Class Activation Mapping, was largely employed as a means of explaining inference. GradCAM can visualize which areas of an image were important for a particular prediction made by a convolutional neural network (CNN).

Introduced in a seminal paper [13], the idea behind GradCAM is to understand which regions of the input image were crucial for a model to make its prediction. It works by computing the gradient of the model output with respect to the feature maps in a particular CNN layer. These gradients represent the importance of each feature map for the final prediction. The feature maps are then weighted by their gradients and the weighted maps are summed up to produce a heatmap highlighting the important regions in the input image.

By the same token, visual means were used for explainability in studies that dealt with timeseries data [14-16]. For instance, one study [14] utilized heatmaps as a visual representation to emphasize the crucial variables across the temporal sequence. Additionally, the method incorporated a channel-wise CNN design to allow for the individual consideration of the timeseries variables. The approach was applied to predict the risk of in-hospital mortality. Another study [16] presented a technique called ECGradCAM, which was used for interpreting deep learning models employed in electrocardiogram (ECG) analysis. The method generates attention maps to explain the decision-making process behind Deep Learning predictions.

In the NLP domain, transformers models, particularly pre-trained transformers, are becoming increasingly popular in healthcare for their ability to handle large amounts of complex data for a variety of NLP tasks. For example, transformer models have been used for sentiment analysis of drug reviews [2], information extraction from clinical text [17], and disease classification based on symptoms described in patient narratives [18, 19]. Additionally, transformers models can be fine-tuned for specific healthcare tasks, such as predicting the risk of adverse events [20], predicting disease progression [21], or generating personalized treatment plans [22].

However, the literature is lacking in studies that have applied XAI for transformer models in the healthcare domain. Despite the growing interest in XAI and its potential benefits for improving the transparency and interpretability of AI models, the healthcare sector has yet to fully embrace this approach. This should be surprising given the importance of transparency and trust in the medical domain, especially when it comes to decisions that impact the health and well-being of patients. In this respect, this study offers a practical example of using explainable NLP in emergency departments. Our goal is to predict patient hospitalization based on triage notes. Equally important, we endeavor to make the model's predictions transparent by identifying the keywords or phrases in the input text that have an impact on the prediction.

III. METHODOLOGY

A. Dataset

The dataset used in this study was sourced from the Amiens-Picardy University Hospital in France. At the outset, it was aimed to comply with ethical and legal regulations. The approval of treatment was obtained under the reference number PI2019_843_0066 from the “Hors Loi Jarde Committee”.

The data consisted of over 260K ED records covering a period of over four years, from January 2015 to June 2019. Every record was linked to a binary label indicating the result of triage, either hospitalization or discharge. The mean length of stay (LOS) spent in the ED unit was roughly 4 hours and 23 minutes, while the average LOS for discharged patients was approximately 4 hours and 4 minutes. The ratio of hospitalization over discharge was about 35%.

The dataset was comprised of a mixture of numeric and categorical variables, as well as textual notes. Numeric variables mostly depicted patients' vital signs such as temperature, heart rate, etc. Categorical variables, on the other hand, provided general information about patients like their gender, origin, family status, etc. A set of four textual fields contained the triage observations, surgical history, psychiatric history, and medical history of patients. All these free-text notes were recorded in French by nurses or physicians during the ED triage. Our attention was solely on these textual fields, which will be discussed in the following sections.

Additionally, the data records specified the specialty that was assigned to hospitalized patients. The majority of the specialties could be grouped into three main categories, including short-term hospitalization, surgery, and medical specialties such as cardiology and neurology. Table 1 provides the statistics of specialties in the dataset.

TABLE 1. SUMMARY OF MEDICAL SPECIALTIES.

Specialty / Label	Hospitalization %
Surgery	19.7%
Short-Term Hospitalization Unit	42.4%
Medical Specialty	33%
Other	4.9%

B. Data Preprocessing

In this study, the emphasis was exclusively placed on the text notes for the development of our models. As such, all numerical and categorical variables were excluded from the dataset at the outset. The exclusion of numerical and categorical variables allowed to simplify the data and focus on the language aspect of the triage notes, thus helping to achieve the aim of creating explainable NLP models.

As mentioned earlier, the dataset contained four columns, which stored the triage notes. Initially, these columns were combined into a single field, creating a unique document for each record. However, it is worth noting that the dataset had numerous blank text fields across thousands of records, with approximately 70K records lacking any textual information. That said, we had to exclude the records that did not contain any text data at all. Further, we excluded the records whose text was shorter than 25 characters, as we were interested in triage notes that contained sufficient information for understanding the reasoning behind predictions. Eventually, the dataset included about 162K records.

Subsequently, standard procedures of text normalization were applied to clean and standardize the textual data. The initial procedures included case conversion, removing stop words and special characters. The normalization process was implemented using Python libraries including NLTK [23].

C. Classification Models

For the classification task, we fine-tuned popular BERT models for the French language. These models were originally pretrained on vast French text datasets, including:

CamemBERT [24]: One of the state-of-the-art language models for the French language developed by the Inria research institute. It is based on the popular BERT [20] architecture and has been pretrained on a large French corpus, making it particularly well-suited for NLP tasks such as text classification, named entity recognition, and question answering. CamemBERT has been utilized in numerous NLP studies dealing with the French language (e.g. [25, 26]).

We employed the camembert-base version, which is one of the six variations available of CamemBERT. Each variant has distinct characteristics such as the number of parameters, amount of data utilized during the pretraining phase, and the domains from which the pretraining was sourced.

FlauBERT [27]: FlauBERT is a variation of the BERT model that incorporates a deep, bidirectional attention mechanism to comprehend the context within a sentence. It was specifically built for the French language and has been trained on a corpus of French text. As a result, FlauBERT delivers state-of-the-art performance on various tasks [28], and is gaining popularity within the French NLP community.

The pretrained transformer models were obtained from the HuggingFace repository [29], which is a widely utilized for the distribution of such models. The ktrain library [30] was employed, which acts as a convenient wrapper for the Keras [31] Deep Learning library. ktrain simplifies the process of training and deploying models through its user-friendly and intuitive API, and it includes built-in support for popular models like BERT [7] and GPT-2 [32], among others.

D. Model Explainability

We utilized the LIME library [33] for providing the model explainability. LIME works by approximating the model's decision boundary around a specific prediction and then fitting a simple, interpretable model to this boundary. This simple model can then be used to understand which input features had the most impact on the prediction. LIME is model-agnostic, meaning it can be used to explain the predictions of any ML model, regardless of the type of model or algorithm used.

In the context of NLP, LIME can be used to explain why a particular text was classified in a certain way by the model. This is done by generating an interpretable representation of the input text that highlights the most important words or phrases that contribute to the prediction. LIME can be especially useful with complex NLP models, as it provides a way to understand the underlying decision-making process, and to identify potential biases. Additionally, because LIME is model-agnostic, it can be used with a wide range of NLP models, including Deep Learning models like transformers.

IV. Experiments

A. Experimental Set-up

The data pre-processing and anonymization procedures were carried out within the hospital's computing facilities to safeguard the confidentiality of patient information. Subsequently, the ML experiments were implemented on the MatriCS platform hosted by the University of Picardie Jules Verne [34]. MatriCS provides advanced computing resources, including high-end GPUs. In particular, the models were fine-tuned on a bi-GPU node featuring two V-100 Nvidia GPUs.

The dataset was divided into train and test sets through a 3-fold cross-validation process, with 20% of the train data for validation. The splitting was performed using the Scikit-Learn library [35].

B. Model Fine-Tuning

The fine-tuning experiments of transformers and their corresponding runtimes are summarized in Table 2. The CamemBERT and FlauBERT models were capable of achieving comparable performance results. Table 3 provides a summary of the performance metrics of both models.

TABLE 2. SUMMARY OF EXPERIMENTAL MODELS.

Model	Params	Embedding Dimension	Runtime
CamemBERT	110M	768	13.3 hrs
FlauBERT	137 M	768	13.5 hrs

TABLE 3. MODEL PERFORMANCE.

Model	AUROC	Precision	Recall	F1-Score
CamemBERT	0.75	0.69	0.70	0.70
FlauBERT	0.74	0.69	0.69	0.69

V. Discussion and Analysis

In this section, we examine the explainability of predictions. We focus solely on the CamemBERT model since it generally achieved better performance.

To assess the explainability of the model, we applied LIME to a subset of texts from the test dataset and carefully scrutinized the explanations provided. To ensure the medical relevance and accuracy of the explanations, the evaluation was conducted in collaboration with a physician who possesses domain expertise in emergency medicine and is one of the co-authors of this paper. This collaboration allowed for insightful interpretations of the model's predictions, aligning them with the medical knowledge and expertise of our team.

We provide a couple of examples below, where the triage notes have been translated into English for better understanding. Figure 1 showcases a case of a discharged patient predicted by the CamemBERT model. On the other hand, Figure 2 presents a case of a patient requiring hospitalization. These examples serve to demonstrate the predictive capabilities of the model for different admission outcomes, highlighting its potential utility in the ED settings.

The visualization provides insights into the significant role played by different words in the prediction. In the visualization, words are color-coded to indicate their impact on the prediction outcome. Words with green shading indicate a positive effect on the prediction, while those with red shading suggest a negative effect. The intensity of the color reflects the magnitude of the coefficients in the linear model derived by LIME. This visualization aids in understanding the contribution of individual words towards the model's prediction.

Effort dyspnea throat constriction no sign of respiratory complication describes palpitations EKG done and checked intermittent palpitations progress since ten days low right chest pain spontaneous resolving recent biology DDM neg.

Figure 1. Example of explaining the prediction of a discharged patient (i.e. True Negative).

Adenocarcinoma of the right colic angle operated on February pt r contiguous pancreatitic involvement liver metastases treated by cephalic duopancratectomy right hemicolectomy ileostomy cholecystectomy hepatic wedge segment vii adjuvant treatment with folfox courses mai December weaned smoking since one year oh cirrhosis child B left wrist fracture years amputation fingers iii iv v right hand work accident anorexia since week vomits colicneoplasia evolving w/ meta

Figure 2. Example of explaining the prediction of an admitted patient (i.e. True Positive), probability of admission ≈64%.

In Figure 1, from a medical perspective, the presence of effort dyspnea supports the prediction of discharge since it is not a permanent symptom and indicates that the patient can continue to be treated by their cardiologist in an outpatient setting. Conversely, respiratory complications suggest a more critical condition that may require hospitalization. Palpitations are often managed through ambulatory care pathways. The presence of chest pain can be associated with both critical conditions (e.g., heart attack) and non-critical causes (e.g., musculoskeletal pain). The words "resolving" and "negative" have been identified as indicators favoring discharge and are likely the most influential in determining the probability of being discharged. These medical observations provide valuable insights into the factors considered by the model for predicting patient admission outcomes.

In Figure 2, patients diagnosed with alcoholic cirrhosis, particularly those with liver metastases, can rapidly deteriorate and require critical care. Conversely, a patient who has successfully quit smoking is more likely to be considered for non-admission. Although a wrist fracture typically requires immobilization with a cast rather than surgery, the presence of a surgical history, even if unrelated to the current reason for seeking care, can contribute to the decision for discharge. Interestingly, it is surprising that duopancreatectomy (a complex surgery involving the pancreas and duodenum) has the highest influence favoring discharge, even though it would generally be expected to lean towards admission. These explanations shed light on the factors contributing to the relatively low admission probability of 64%.

VI. Conclusions

In conclusion, our study highlights the importance of the triage model in assisting with patient admission decisions, while acknowledging that it does not possess all the comprehensive medical data. Interestingly, there were instances where, as physicians, we would have made similar decisions as the model based on the available data. The NLP-based explanations could offer valuable insights for physicians, aiding in the understanding of the model's high or low probabilities of admission. However, it is crucial to note that certain medical concepts considered as pro-admission by the model may differ from a physician's perspective. Moving forward, the integration of a human-in-the-loop approach in a larger study holds promise for enhancing the model by leveraging the medical review of explanations. This collaborative effort between AI and medical professionals has the potential to refine and improve patient admission predictions in emergency medicine.

Identify applicable funding agency here. If none, delete this text box.

References

[1] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." *In Proceedings of Advances in Neural Information Processing Systems (NIPS)* 30 (2017).

[2] Punith, N. S., and Krishna Raketla. "Sentiment analysis of drug reviews using transfer learning." *In Proceedings of the 3rd International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 1794-1799. IEEE, 2021.

[3] Yang, Jiacheng, Mingxuan Wang, Hao Zhou, Chengqi Zhao, Weinan Zhang, Yong Yu, and Lei Li. "Towards making the most of BERT in neural machine translation." *In Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, pp. 9378-9385. 2020.

[4] He, Yun, Ziwei Zhu, Yin Zhang, Qin Chen, and James Caverlee. "Infusing disease knowledge into BERT for health question answering, medical inference and disease name recognition." *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* .2020.

[5] Mahmoud Elbattah, Arnaud, Émilien, , Maxime Gignon, and Gilles Dequen. "The Role of Text Analytics in Healthcare: A Review of Recent Developments and Applications". *In Proceedings of the 14th International Joint Conf. on Biomedical Engineering Systems and Technologies (BIOSTEC)*. 2021.

[6] Rothstein, Mark A. "Ethical issues in big data health research." *Journal of Law, Medicine & Ethics* 43, no. 2 (2015): 425-429.

[7] Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-training of deep bidirectional transformers for language understanding." *In Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*. 2019.

[8] Arnaud, Émilien, Mahmoud Elbattah, Maxime Gignon, and Gilles Dequen. "Deep learning to predict hospitalization at triage: Integration of structured data and unstructured text." *In Proceedings of the IEEE International Conference on Big Data*, pp. 4836-4841. IEEE, 2020.

[9] Corizzo, Roberto, Yohan Dauphin, Colin Bellinger, Eftim Zdravevski, and Nathalie Japkowicz. "Explainable image analysis for decision support in medical healthcare." *In Proceedings of the IEEE International Conference on Big Data*, pp. 4667-4674. IEEE, 2021.

[10] Basu, Sanhita, Sushmita Mitra, and Nilanjan Saha. "Deep learning for screening covid-19 using chest x-ray images." *In Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 2521-2527. IEEE, 2020.

[11] Oh, Yujin, Sangjoon Park, and Jong Chul Ye. "Deep learning COVID-19 features on CXR using limited training data sets." *IEEE Transactions on Medical Imaging* 39, no. 8 (2020): 2688-2700.

[12] Prabhushankar, Mohit, and Ghassan AlRegib. "Extracting causal visual features for limited label classification." *In Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 3697-3701. IEEE, 2021.

[13] Selvaraju, Ramprasaath R., Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. "Grad-CAM:

Visual explanations from deep networks via gradient-based localization." *In Proceedings of the IEEE International Conference on Computer Vision*, pp. 618-626. 2017.

[14] Viton, Fabien, Mahmoud Elbattah, Jean-Luc Guérin, and Gilles Dequen. "Heatmaps for visual explainability of CNN-based predictions for multivariate time series with application to healthcare." *In Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 1-8. IEEE, 2020.

[15] Wickstrøm, Kristoffer, Karl Øyvind Mikalsen, Michael Kampffmeyer, Arthur Revhaug, and Robert Jenssen. "Uncertainty-aware deep ensembles for reliable and explainable predictions of clinical time series." *IEEE Journal of Biomedical and Health Informatics* 25, no. 7 (2020): 2435-2444.

[16] Hicks, Steven A., Jonas L. Isaksen, Vajira Thambawita, Jonas Ghouse, Gustav Ahlberg, Allan Linneberg, Niels Grarup et al. "Explaining deep neural networks for knowledge discovery in electrocardiogram analysis." *Scientific Reports* 11, no. 1 (2021): 10949.

[17] Arnaud, Émilien, Mahmoud Elbattah, Maxime Gignon, and Gilles Dequen. "Learning Embeddings from Free-text Triage Notes using Pretrained Transformer Models." *In Proceedings of 15th International Conference on Health Informatics (HEALTHINF)*, pp. 835-841. 2022.

[18] Li, Xuedong, Walter Yuan, Dezhong Peng, Qiaozhu Mei, and Yue Wang. "When BERT meets Bilbo: a learning curve analysis of pretrained language model on disease classification." *BMC Medical Informatics and Decision Making* 21, no. 9 (2021): 1-11.

[19] Çelıkten, Azer, and Hasan Bulut. "Turkish medical text classification using BERT." *In Proceedings of the 29th Signal Processing and Communications Applications Conference (SIU)*, pp. 1-4. IEEE, 2021.

[20] Jarynowski, Andrzej, Alexander Semenov, Mikołaj Kamiński, and Vitaly Belik. "Mild adverse events of sputnik v vaccine extracted from russian language telegram posts via BERT deep learning model." *Journal of Medical Internet Research*, Nov 29;23(11) (2021): 2021-06.

[21] Hu, Wendeng, and Sophia Y. Wang. "Predicting glaucoma progression requiring surgery using clinical free-text notes and transfer learning with transformers." *Translational Vision Science & Technology* 11, no. 3 (2022): 37-37.

[22] Vo, Chau, and Tru Cao. "Association Discovery from Electronic Medical Records towards Personalized Treatment." *In Proceedings of the 9th NAFOSTED Conference on Information and Computer Science (NICS)*, pp. 134-141. IEEE, 2022.

[23] Bird, Steven, Ewan Klein, and Edward Loper. Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc.", 2009.

[24] Martin, Louis, Benjamin Muller, Pedro Javier Ortiz Suárez, Yoann Dupont, Laurent Romary, Éric Villemonte de La Clergerie, Djamé Seddah, and Benoît Sagot. "CamemBERT: a tasty French language model." *arXiv preprint arXiv*:1911.03894 (2019).

[25] Sauvayre, Romy, Jessica Vernier, and Cédric Chauvière. "An Analysis of French-Language Tweets About COVID-19 Vaccines: Supervised Learning Approach." *JMIR Medical Informatics* 10, no. 5 (2022).

[26] Martin, Guillaume L., Julien Jouganous, Romain Savidan, Axel Bellec, Clément Goehrs, Mehdi Benkebil, Ghada Miremont et al. "Validation of artificial intelligence to support the automatic coding of patient adverse drug reaction reports, using nationwide pharmacovigilance data." *Drug Safety* 45, no. 5 (2022): 535-548.

[27] Le, Hang, Loïc Vial, Jibril Frej, Vincent Segonne, Maximin Coavoux, Benjamin Lecouteux, Alexandre Allauzen, Benoit Crabbé, Laurent Besacier, and Didier Schwab. "Flaubert: Unsupervised language model pre-training for French." *arXiv preprint arXiv*:1912.05372 (2019).

[28] Blanc, Corentin, Alexandre Bailly, Élie Francis, Thierry Guillotin, Fadi Jamal, Béchara Wakim, and Pascal Roy. "FlauBERT vs. CamemBERT: Understanding patient's answers by a French medical chatbot." *Artificial Intelligence in Medicine* 127 (2022): 102264.

[29] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2019). Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv*:1910.03771.

[30] Maiya, Arun S. "ktrain: A low-code library for augmented machine learning." *Journal of Machine Learning Research* 23 (2020): 1-6.

[31] Chollet, F. (2015). Keras. GitHub Repository: https://github.com/fchollet/keras.

[32] Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. "Language models are unsupervised multitask learners." *In Proceedings of the International Conference on Learning Representations (ICLR)*. 2019.

[33] Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. "" Why should I trust you?" Explaining the predictions of any classifier." *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135-1144. 2016.

[34] MatriCS Computing Platform: https://www.matrics.u-picardie.fr/en/home/description/

[35] Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel et al. "Scikit-learn: Machine learning in Python." *Journal of Machine Learning Research* 12 (2011): 2825-2830.